Science and Technology for the Built Environment

ISSN: 2374-4731 (Print) 2374-474X (Online) Journal homepage: www.tandfonline.com/journals/uhvc21

Model-free optimal chiller loading method based on Q-learning

Shunian Qiu, Zhenhai Li, Zhengwei Li & Xinfang Zhang

To cite this article: Shunian Qiu, Zhenhai Li, Zhengwei Li & Xinfang Zhang (2020) Model-free optimal chiller loading method based on Q-learning, Science and Technology for the Built Environment, 26:8, 1100-1116, DOI: 10.1080/23744731.2020.1757328

To link to this article: https://doi.org/10.1080/23744731.2020.1757328

Published online: 13 May 2020.

Science and Technology for the Built Environment, (2020) **26**, 1100–1116

ISSN: 2374-4731 print / 2374-474X online
DOI: 10.1080/23744731.2020.1757328

Model-free optimal chiller loading method based on Q-learning

SHUNIAN QIU[1], ZHENHAI LI[1], ZHENGWEI LI[1,2]* and XINFANG ZHANG[1]

[1]*School of Mechanical and Energy Engineering, Tongji University, Shanghai, China*
[2]*Key Laboratory of Performance Evolution and Control for Engineering Structures of Ministry of Education, Tongji University, Shanghai, China*

Chillers consume considerable energy in building HVAC systems, and the optimal operation of chillers is essential for energy conservation in buildings. This article proposes a model-free optimal chiller loading (OCL) method for optimizing chiller operation. Unlike model-based OCL methods, the proposed method does not require accurate chiller performance models as a priori knowledge. The proposed method is based on the Q-learning method, a classical reinforcement learning method. With the comprehensive coefficient of performance (COP) of chillers as the environmental feedback, the model-free loading controller can learn autonomously and optimize the chiller loading by adjusting the set points of the chilled water outlet temperature. A central chiller plant in an office building located in Shanghai is selected as a case system to investigate the energy conservation performance of the proposed method through simulations. The simulation results suggest that the proposed method can save 4.36% of chiller energy during the first cooling season compared to the baseline control, which is slightly inferior to the value for the model-based loading method (4.95%). Owing to its acceptable energy-saving capability, the proposed method can be applied to central chiller plants that lack a system model and historical data.

Introduction

Optimal chiller loading (OCL)

More than half of a building's total energy consumption is for the HVAC system (Pérez-Lombard, Ortiz, and Pout 2008; Li et al. 2016; Hou et al. 2018; Pang et al. 2018, 2020). The chiller is a typical item of equipment for the HVAC system (Kreider 1994; Qiu et al. 2018). According to statistical data (Commercial Buildings Energy Consumption Survey (CBECS) 2012) from a commercial buildings energy consumption survey (CBECS), central chillers are widely used in large buildings. Among buildings with building area over 200,000 square feet (18,580 m^2), 50% are equipped with central chillers; among buildings with 10 or more floors, 61% are equipped with central chillers (CBECS 2012).

Chillers consume considerable energy (Taylor 2012; Ardakani, Ardakani, and Hosseinian 2008; Chang, Lin, and Chuang 2005). The chiller operation must be optimized to reduce the chiller energy consumption (Taylor 2012). Typically, the chiller optimization can be classified into two categories: optimal chiller sequencing (OCS) and optimal chiller loading (OCL) (Chang 2004; Beghi, Cecchinato, and Rampazzo 2011). OCS is intended to reduce the chiller energy by reducing unnecessary chiller operations through optimizing the sequence of the chiller on–off status (Chang, Lin, and Lin 2005; Li, Huang, and Sun 2014; Sun, Wang, and Huang 2009). OCL methods are used to enhance the chiller coefficient of performance (COP) by optimizing the distribution of the total system cooling load to multiple chillers because the chiller COP is highly correlated with the partial load ratio (PLR) of the chiller (Chang, Lin, and Chuang 2005; Lee and Lin 2009; Qiu, Feng, Zhang, et al. 2019).

OCL methods have been investigated by many researchers. Among the OCL methods developed, model-based methods, which are also termed "model-based supervisory control," are studied most commonly (Wang and Ma 2008). A typical workflow of model-based OCL methods is illustrated in Figure 1. They are usually composed of (1) a chiller performance model, which is determined by the historical operational data or equipment manual of the target system; (2) an optimization algorithm, such as a genetic algorithm (GA) or particle swarm optimization (PSO); (3) real-time environmental data (e.g., system cooling load data); and (4) an optimization objective function. The performance model and optimization objective function are for the evaluation of the loading plans, and the optimization algorithm(s) is (are) used to determine the optimal loading plan. This workflow profiles the following model-based OCL studies.

Lee and Lin (2009) developed an OCL method with the PSO as an optimization technique, a power–PLR model as the chiller performance model, and system cooling load as the environmental variable input. In their method, the supplied chilled water temperature set points ($T_{chws,set}$) are used

Received December 31, 2019; Accepted April 14, 2020
Shunian Qiu is a Student, **Zhenhai Li**, **PhD**, is a Professor, **Zhengwei Li**, **PhD**, is an Associate Professor, **Xinfang Zhang** is a Student.
*Corresponding author e-mail: zhengwei_li@tongji.edu.cn

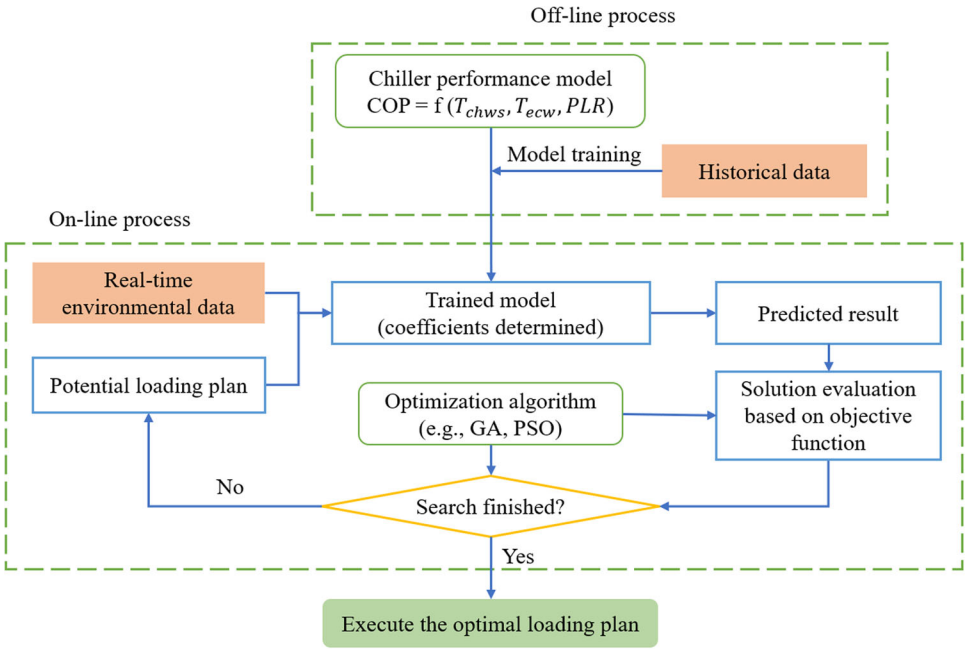

Fig. 1. Typical work flow of model-based OCL methods.

as control actions to distribute the cooling load to the chillers. Several similar model-based OCL methods are listed in Table 1, and the respective variables are identified in the nomenclature section of this article.

Defects of model-based OCL methods

The energy conservation performance of the model-based OCL methods has been investigated and validated. For buildings with accurate chiller models, optimized chiller operation could be realized via mature model-based OCL methods, such as the ones reviewed in Table 1. However, model-based methods still have shortcomings regarding the model.

(1) *High reliability of accurate models*: The "model" is the base and key of model-based optimal control methods. All methods reviewed in the preceding section require accurate chiller performance models. However, a model always contains uncertainties owing to the simplified model structure and error in the model parameters (Zhu, Shan, et al. 2013; Pang and O'Neill 2018). Inaccurate models may result in wrong control decisions, which can affect the system operation performance.

(2) *Accurate models are difficult to acquire*: In published studies, the chiller models are obtained either from the equipment manual (Chang, Lin, and Chuang 2005; Beghi, Cecchinato, and Rampazzo 2011; Lee and Lin 2009) or from regression of the historical operational data (Swider 2003; Wang et al. 2018). Simple chiller models such as the COP–PLR model are easy to obtain from equipment manuals. However, this model type does not consider the influence of other variables on the chiller performance, which can affect the model accuracy. Multivariate models (e.g., the multivariate polynomial and neural network models; Swider 2003) usually exhibit better accuracy. However, they require sufficient historical operational data to determine a number of coefficients correctly. Also, the historical operational database takes time to accumulate.

In brief, model-based OCL methods require accurate chiller performance models to achieve good energy conservation, while an accurate chiller model is difficult to acquire. For those buildings without accurate chiller models or sufficient data (yet), the optimization of chiller operation remains a challenge.

Reinforcement learning and its application in building control

The machine learning field can be classified into three subdomains: supervised, unsupervised, and reinforcement learning (RL). Supervised machine learning is intended to mine the correlation between the feature matrix X and label array y (Fulkerson 1995). Unsupervised learning recognizes patterns in datasets that cannot be manually labeled because of a lack of existing knowledge (Wang and Xiao 2004; Abdi and Williams 2010; Arthur and Vassilvitskii 2007).

RL is more dynamic than the other two. Generally, RL is intended to design a learning agent to play a game (e.g., chess) (Sutton, Barto, and Bach 2018). In each step of the game, the agent must decide the next move, make the move, obtain the feedback from the environment, and update its own experience. RL algorithms are used to guide the agent to build/update/use its own experience. Figure 2 shows a simple instance based on the Q-learning method (a classical RL method) (Sutton, Barto, and Bach 2018).

Table 1. Typical model-based OCL methods.

Researchers	Required chiller model	Optimization algorithm	Optimization objective	Real-time environmental data	Energy-saving rate of chillers
Lee and Lin (2009)	$P_{chiller} = f(PLR, PLR^2, PLR^3)$	PSO	Minimum total power of chillers $\mathrm{Min}\sum_i P_{chiller,i}$	System cooling load CL_s	0–9.9% (compared to Lagrangian method) 0.2–10.3% (compared to GA)
Chang, Lin, and Lin (2005)	$P_{chiller} = f(PLR, PLR^2, PLR^3)$	GA	$\mathrm{Min}\sum_i P_{chiller,i}$	CL_s	0.1–0.7% (compared to Lagrangian method)
Ardakani, Ardakani, and Hosseinian (2008)	$P_{chiller} = f(PLR, PLR^2, PLR^3)$	PSO and GA	$\mathrm{Min}\sum_i P_{chiller,i}$	CL_s	0.1–1.4% (compared to binary GA)
Coelho et al. (2014)	$P_{chiller} = f(PLR, PLR^2, PLR^3)$ and $P_{chiller} = f(PLR, PLR^2)$	New differential cuckoo search approach	$\mathrm{Min}\sum_i P_{chiller,i}$	CL_s	0–3.7% (compared to PSO) 0–3.3% (compared to evolution strategy)
Beghi, Cecchinato, and Rampazzo (2011)	$P_{chiller} = f(T_{chwr}, T_{air}, F_{chw}, PLR)$ $CC = f(T_{chwr}, T_{air}, F_{chw}, PLR)$	Multiphase GA (MPGA)	Minimum total energy consumption of chillers $\mathrm{Min}\sum_i Energy_i$	$T_{chwr}, T_{air}, F_{chw}, CL_s$ (for air-cooled chillers)	2.5% (compared to symmetric strategy) 1.3% (compared to sequential strategy)
Chang (2004)	$COP = f(PLR, PLR^2)$	Lagrange multiplier	Maximum sum of chiller COPs $\mathrm{Max}\sum_i COP_i$	CL_s	0.1–2.4% (compared to equal loading)

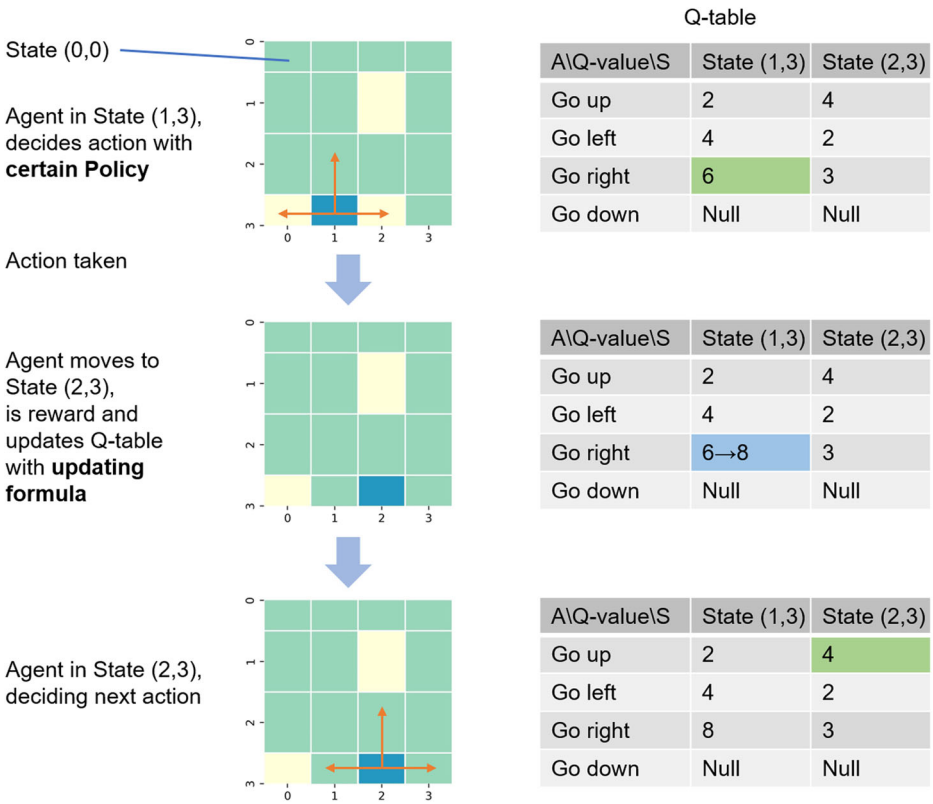

Fig. 2. Simple example of RL application: mouse–maze–cheese game.

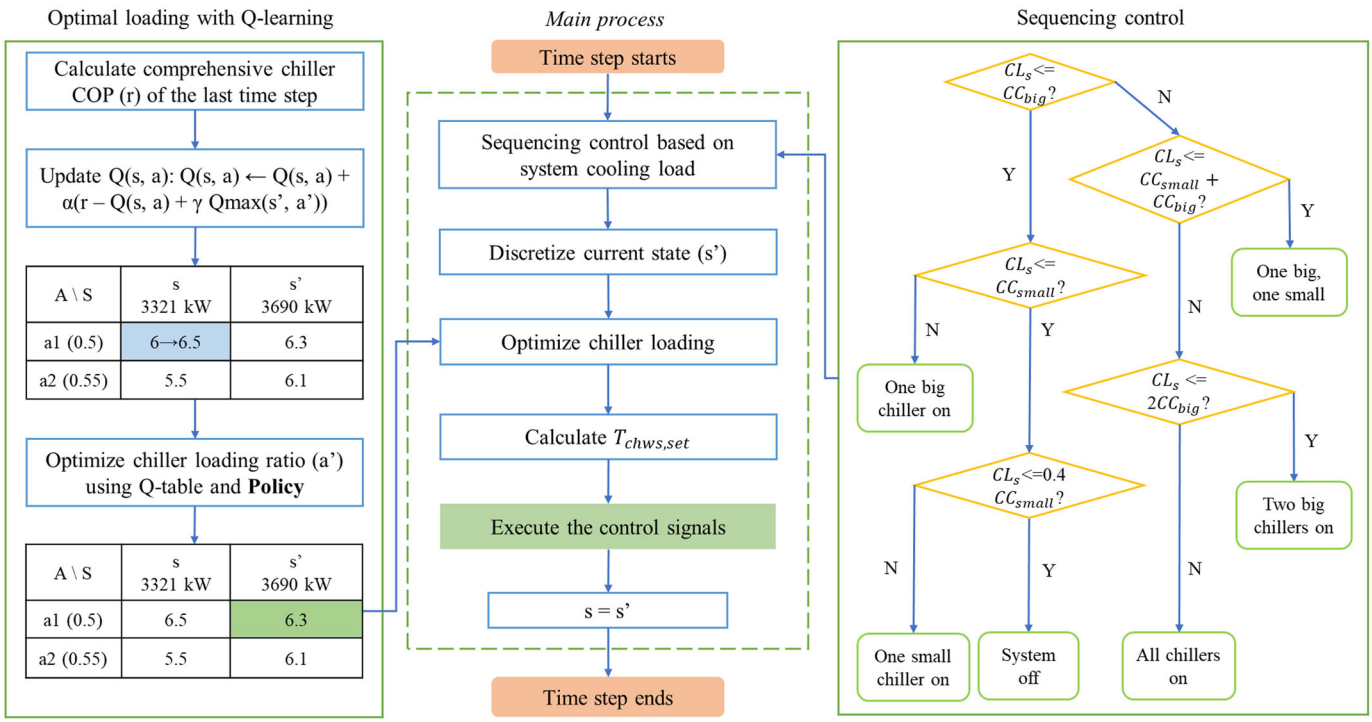

Fig. 3. Work flow of proposed control method.

In Figure 2, a mouse (agent) travels in a maze searching for cheese (maximal reward), which lies on points (0, 3), (2, 1), and (2, 3) of the maze. When a time step starts, the agent decides the next action (i.e., its movement) based on its experience (Q-table), which is a two-dimensional (2D) table composed of states, actions, and Q-values. The Q-value indicates the worth of an action at a certain state. After executing the selected action, the agent obtains a reward from the environment. Based on the reward, the agent evaluates the worth of the last movement in the last state and updates the Q-table, which is shown in Figure 2 as (6 → 8). Subsequently, the next time step starts.

Applications of RL in the building control domain are mostly for the control of renewable energy equipment. Henze and Schoenmann (2003) adopted Q-learning to optimize the charge–discharge action of energy storage equipment without any prediction or system models. Moreover, they investigated how to choose state variables from the state-of-charge, cooling loads, and real-time pricing rates when the reward variable varies.

To develop an RL agent to control the energy storage equipment, Liu and Henze (2006a, 2006b) applied two steps: (1) pretraining: train the RL agent in a virtual environment; and (2) posttraining: train the agent in the real environment. The agent action space in their study consists of the set point of the global zone air temperature and the charge–discharge rate of the thermal energy storage equipment.

Moreover, Yang et al. (2015) used batch Q-learning with memory replay to optimize the operation of a building equipped with PV/T panels and geothermal heat pumps. In their study, the entire system is decomposed into three loops: (1) PV/T panel–heat exchanger loop, (2) heat exchanger–ground–geothermal heat pump loop, and (3) heat pump–conditioned space loop. The flow rates of the thermal media in each loop are optimized.

Motivation of this research study

Model-based OCL methods have been widely investigated. Although they exhibit a good chiller energy-saving performance, their defects are nonnegligible. Therefore, it is worth investigating how to optimize chiller loading without prepared accurate models or sufficient historical data. To tackle the problem, a model-free OCL method based on RL is developed in this study. The proposed method only requires the system layout and equipment nameplates as the prepared information, which could be acquired simply by checking the equipment room. Hence, this method could offer the buildings with limited a priori knowledge a feasible option to optimize the chiller operation.

The methodology section of this article presents the proposed method, and the case study section introduces a simulation case study that compares the energy conservation performance of the model-based chiller loading method and proposed model-free method. The simulation results are presented and analyzed in the results and discussion section, with the conclusion and description of future work wrapping up the article.

Methodology

Overview

First, the important features and hypothesis of the proposed method are addressed as follows.

Limitation: The proposed method is designed to optimize the chiller loading of central chiller plants composed of two sizes of chillers (big/small). Moreover, the proposed method assumes that the performances of equally sized chillers are equal. In other words, if a central chiller plant includes n identical small chillers and m identical big chillers, this system is suitable for the proposed method. In addition, the targeted system of this method is a decoupled system, the chilled water loop of which is composed of primary side and secondary side. The chilled water flow rate of each chiller is constant, while the flow rate on the secondary side is adjustable (Chang, Lin, and Lin 2005; Gao et al. 2016).

Simplification: Based on the limitation of the method, the OCL problem can be simplified based on how to optimize the cooling load distribution between the big and small chiller(s). The optimal loading among identical chillers is equal loading according to the calculation process developed by Chang (2004).

Optimization objective: The optimization objective in this study is the comprehensive COP of all operating chillers, which is calculated with Equation 1. The controlled variables are set points of the chilled water temperature that exits the chillers ($T_{chws,set}$):

$$Comprehensive\ chiller\ COP = \frac{CL_s}{\sum P_{chiller,i}} \quad (1)$$

where CL_s is the system cooling load (kW) and $P_{chiller,i}$ is the power of the ith chiller (kW).

Required a priori knowledge and real-time data: Before the application, the proposed method requires certain prepared information of the system layout and equipment name plates. During the system operation, the proposed method requires real-time values of the system cooling load (Equation 2) and the chiller power. Compared with the model-based control method, the proposed method does not require chiller performance models:

$$CL_s = C_p \times \rho \times F_{chw} \times (T_{chwr} - T_{chws}) \div 3600 s/h \quad (2)$$

where T_{chws} is the temperature of the chilled water in the header supply pipe (°C), T_{chwr} the temperature of the chilled water in the header return pipe (°C), F_{chw} the total chilled water flow rate (m^3/h), C_p the specific heat capacity of water = 4.2 kJ/(kg·K), and ρ the water density = 1000 kg/m^3.

Optimization algorithm: This is Q-learning of RL. As illustrated in Figure 3, the proposed method is realized by the following steps:

A. When a time step starts, the proper on–off status of each chiller plant is determined with the sequencing control method based on the system cooling load (conventional method), which is presented in the next seubsection on sequencing control of chiller plants..
B. The current environmental data (i.e., system cooling load) of the system are discretized into the state (s') within a predefined state space. This step is elaborated in a later subsection on definition and initialization of the Q-table .
C. The loading distribution between big and small chiller(s) is optimized with the Q-learning method: (1) Determine the maximal Q-value corresponding to the current state (s') in the Q-table, (2) update the Q-value of the last state (s) with the updating formula (Equation 3) and last reward (i.e., the comprehensive chiller COP), and (3) determine the optimal loading action (i.e., the ratio of the system cooling load assigned to small chiller(s)) with a certain policy. Details are provided in the subsection on the optimization policy and Q-table updating formula.
D. Convert the optimal loading action into chiller control action ($T_{chws,set}$) with Equation 5.
E. Execute all determined control signals for the system.
F. Record the current state (s') as s.

Sequencing control of chiller plants

(1) To protect chillers, the entire central chiller plant remains turned off if the system cooling load is below 40% of the nominal cooling capacity of the smallest chiller (Chang 2004).
(2) The on–off status of the chiller is determined based on the real-time system cooling load to minimize the unnecessary system cooling capacity. The right side in Figure 3 shows the sequencing control logic of the case-study system investigated in this study (one small and two big chillers). For instance, when the system cooling load is below the cooling capacity of the small chiller, only the small chiller is turned on. When the system cooling load is higher than the cooling capacity of one big chiller and below the total cooling capacity of one big and one small chiller, the controller certainly turns on one big and one small chiller (Li, Huang, and Sun 2014; Qiu, Feng, Li, et al. 2019; Liao, Sun, and Huang 2015).
(3) The on–off statuses of the cooling water pumps, primary chilled water pumps, and cooling towers are coordinated to the status of the chillers. For example, when chiller 1 is running, cooling water pump 1, primary chilled water pump 1, and cooling tower 1 are running (Qiu, Feng, Zhang, et al. 2019).
(4) The frequencies of the running pumps and tower fans are maintained at the nominal value to provide the maximal heat exchange capacity of the chiller evaporators and condensers. Moreover, this rule can protect chillers from icing and other potential damage (Chang, Lin, and Chuang 2005; ASHRAE 2000).

Definition and initialization of Q-table

In the proposed method, the cooling loads acting on each operating chiller are optimized by an RL agent with the Q-learning method. In this study, the Q-table is normalized and initialized as follows:

(1) *State:* In this study, the state is defined as the discretized system cooling load (CL_s) because the system cooling load is uncontrollable and not affected by the system operation. Moreover, it is an essential

Table 2. Format of Q-table in this study.

A\S	0.05 MCC	0.1 MCC	……	0.95 MCC	1 MCC
0	Q (s_1, a_1)	Q (s_2, a_1)		Q (s_{19}, a_1)	Q (s_{20}, a_1)
0.05	Q (s_1, a_2)	Q (s_2, a_2)		Q (s_{19}, a_2)	Q (s_{20}, a_2)
……					
1	Q (s_1, a_{21})	Q (s_2, a_{21})	……	Q (s_{19}, a_{21})	Q (s_{20}, a_{21})

variable in the operation of the central chiller plant (Chang, Lin, and Chuang 2005; Li, Huang, and Sun 2014; Lee and Lin 2009; Qiu, Feng, Zhang, et al. 2019; Sulaiman et al. 2015; Chen et al. 2014; Liu et al. 2017).

In the proposed method, the system cooling load is discretized according to the maximal cooling capacity (MCC; i.e., the cooling capacity sum of all chillers) of the applied system. The system cooling load should be discretized into 0.05, 0.1, … …, 0.95, 1 MCC; this array is used as the state space (i.e., Q-table columns). The real-time value of the state (discretized system cooling load) is used as the real-time state.

(2) *Reward:* The comprehensive chiller COP calculated with Equation 1 is taken as the reward in this method, which is in accordance with the optimization objective.

(3) *Q-value:* The Q-value represents the worth of an action. The meaning of the Q-value in this method is the comprehensive chiller COP, which is in accordance with the optimization objective. Therefore, the initial Q-values are 15–20% higher than the nominal comprehensive chiller COP, to stimulate the agent exploration and avoid agent trapping in local optima.

(4) *Action:* In this study, the optimized action of the RL agent is the ratio of the system cooling load distributed to the small chiller(s). For instance, for a ratio = 0.3, 30% of the system cooling load is distributed to the operating small chiller(s) and 70% to the big chiller(s). Also, identical chillers are equally loaded because according to Chang (2004), the best loading solution for identical chillers is equal loading, whereas the loading for different chillers should be optimized with algorithms.

In other words, when n small chillers are operating, each operating small chiller is supplied with $\left(\frac{1}{n}\right)\times$ ratio $\times$ CL_s. Similarly, if the number of running big chillers is m, each running big chiller is supplied with $\left(\frac{1}{m}\right)\times$ (1-ratio) $\times$ CL_s. Moreover, the precision of the ratio is 0.05. Thus, the action space can be addressed as (0, 0.05, 0.1, 0.15, …, 0.95, 1).

The previously presented discretization precisions of the states and actions are defined by considering the exploration cost. A higher precision results in a larger Q-table (i.e., larger action space and state space), which can positively affect the control performance of a well-trained model-free controller. However, it also requires longer exploration and controller training periods (Sutton, Barto, and Bach 2018). To balance the training cost and control accuracy, the precision of state space and action space is defined as 0.05 in this study. Table 2 gives an example of the Q-table in this study.

Optimization policy and Q-table updating formula

As explained in the overview section, when a time step begins, the Q-learning agent updates the Q-table first. Equation 3 is the standard updating formula of the Q-learning method (Sutton, Barto, and Bach 2018):

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \quad (3)$$

where $Q(s,a)$ is the Q-value corresponding to the last state (s) and last action (a), α is the learning rate (0.9 in this study to accelerate the agent learning performance), r is the reward resulting from the last action (a), and γ is the weight of the future reward for the decision of the current action. In this study, γ is set to 0.01 because the agent action does not affect the next state. Hence, in every time step, the target of the agent is to maximize the current reward instead of pursuing a long-term profit. In addition, $\max_{a'} Q(s',a')$ is the maximum Q-value in state s' according to the current Q-table. In the proposed method, the Q-table is updated for the entire system life in case that system degrades (Zhu, Shan, et al. 2013).

After updating the Q-table, the agent must decide the next action based on a certain policy. In this study, a novel modified version of the ε-greedy policy (Sutton, Barto, and Bach 2018) is developed to solve the problem such that the balance between the exploration and exploitation is never changed during the game. This novel policy is described in Equation 4:

$$\pi(as) = \begin{cases} \frac{10\times\frac{q}{p}}{1+10\times\frac{q}{p}} + \frac{Q(s,a)}{\sum Q(s,a)} \times \frac{1}{1+10\times\frac{q}{p}} & if\ \ a = arg\,max_a\, Q(s,a) \\ \frac{Q(s,a)}{\sum Q(s,a)} \times \frac{1}{1+10\times\frac{q}{p}} & if\ \ a \neq arg\,max_a\, Q(s,a) \end{cases} \quad (4)$$

where q is the number of passed time steps and p is a predefined parameter representing a milestone of the agent learning process, which should be defined as the length of an entire cooling season; p is 4416 in this study because the length of the case study is 6 months (an entire cooling season), and the optimization time step is 1 hour (4416 = 24 hours $\times$ 184 days). Equation 4 is explained in the following way:

(1) A bonus (bonus $= \frac{10\times\frac{q}{p}}{1+10\times\frac{q}{p}}$) is added to the probability that the optimal action (yet) is selected. As time passes (as q increases), the bonus increases, and the agent increasingly prefers exploitation to exploration.

(2) At the beginning of the system operation, the action space is not fully explored yet. In this period, exploration is encouraged. The probability of an action being selected is approximately proportional to its Q-value because the bonus is close to zero.

Table 3. Characteristics of the central chiller plant.

Equipment	Number	Nominal characteristics
Centrifugal chiller	2 big, 1 small	Big: power = 504 kW, cooling capacity = 2810 kW, T_{chws} = 6°C, T_{chwr} = 12°C Small: power = 314 kW, cooling capacity = 1760 kW, T_{chws} = 6°C, T_{chwr} = 12°C
Primary chilled water pump	2 big, 1 small	Big: power = 30 kW, flow rate = 403 m^3/h, head = 19 m, constant speed Small: power = 15 kW, flow rate = 252 m^3/h, head = 15 m, constant speed
Cooling water pump	2 big, 1 small	Big: power = 75 kW, flow rate = 581 m^3/h, head = 34 m, constant speed Small: power = 55 kW, flow rate = 366 m^3/h, head = 33 m, constant speed
Cooling tower	2 big, 1 small	Big: power = 30 kW, flow rate = 569 m^3/h, constant speed Small: power = 16.5 kW, flow rate = 392 m^3/h, constant speed

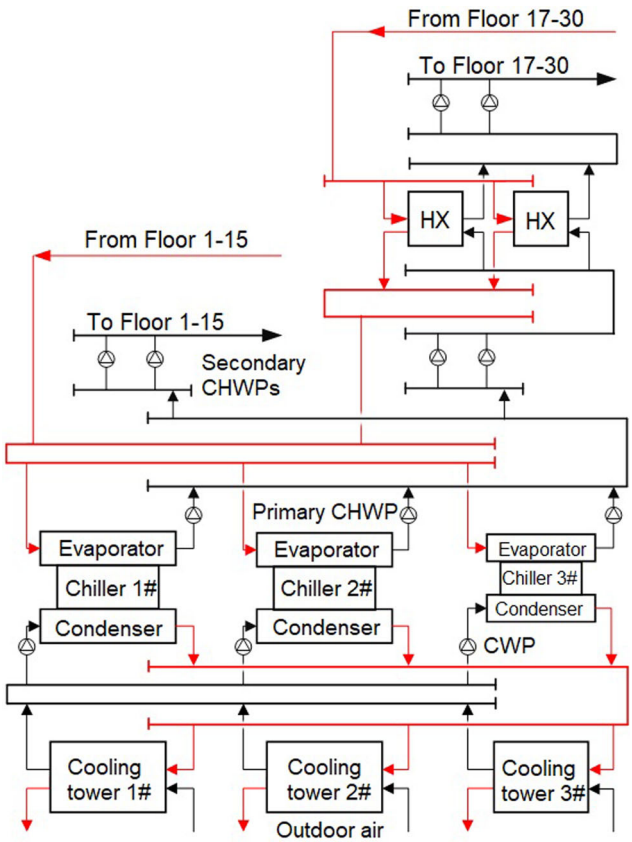

Fig. 4. Layout of the case system.

(3) Although the bonus increases continuously with time, the agent will not completely abandon exploring, no matter how large the bonus becomes.

Expert knowledge interpretation

To avoid unreasonable actions of the agent, expert knowledge should be adopted to interpret the RL agent decision. In this study, the expert knowledge is used for the determination of the control signals and the limitation of the agent action.

(1) As demonstrated in the subsections on definition and initialization of the Q-table and the optimization policy and Q-table updating formula, the designed RL agent can optimize the loading of big and small chiller(s) by optimizing the ratio of the system cooling load acting on the small chiller(s); this action cannot be directly applied to control the chillers. In this method, the chiller loading action is realized by changing the $T_{chws,set}$ of each operating chiller. For each running chiller, $T_{chws,set}$ is calculated with Equation 5:

$$T_{chws,j,set} = T_{chwr,ref} - \frac{CL_j}{C_p \times \rho \times F_{chw,j} \div 3600 s/h} \quad (5)$$

where $T_{chws,j,set}$ is the set point of the jth chiller, $T_{chwr,ref}$ is the nominal T_{chwr} value of the targeted system, CL_j is the cooling load assigned to the jth chiller (kW), C_p is the specific heat capacity of water = 4.2 kJ/(kg·K), $F_{chw,j}$ is the chilled water flow rate of the jth chiller (m^3/h), and ρ is the water density = 1000 kg/m^3.

(2) To avoid potential damage to the chillers, the PLR of a chiller should remain within a certain range to protect the hardware. Some previous OCL studies limited the PLR to 0.3–1.0 (Ardakani, Ardakani, and Hosseinian 2008; Chang, Lin, and Chuang 2005; Lee and Lin 2009), and some studies defined the range as 0.5–1.0 (Chang 2004; Chang, Lin, and Lin 2005). In this study, the PLR is set to 0.4–1.0.

Case study

Case system

A central chiller plant in an office building in Shanghai city is adopted as the case system, and the equipment characteristics on the nameplates are listed in Table 3. The structure of the system is illustrated in Figure 4. This chilled water system is a decoupled system with constant flow rate on each running chiller (Gao et al. 2016). The system cooling load data simulated by EnergyPlus (Crawley et al. 2001; Zhu, Hong, et al. 2013; EnergyPlus 2020) and TMY2 weather data are used as the simulation input. The nominal T_{chwr} value of this system is 12 °C, and the nominal T_{chws} is 6 °C.

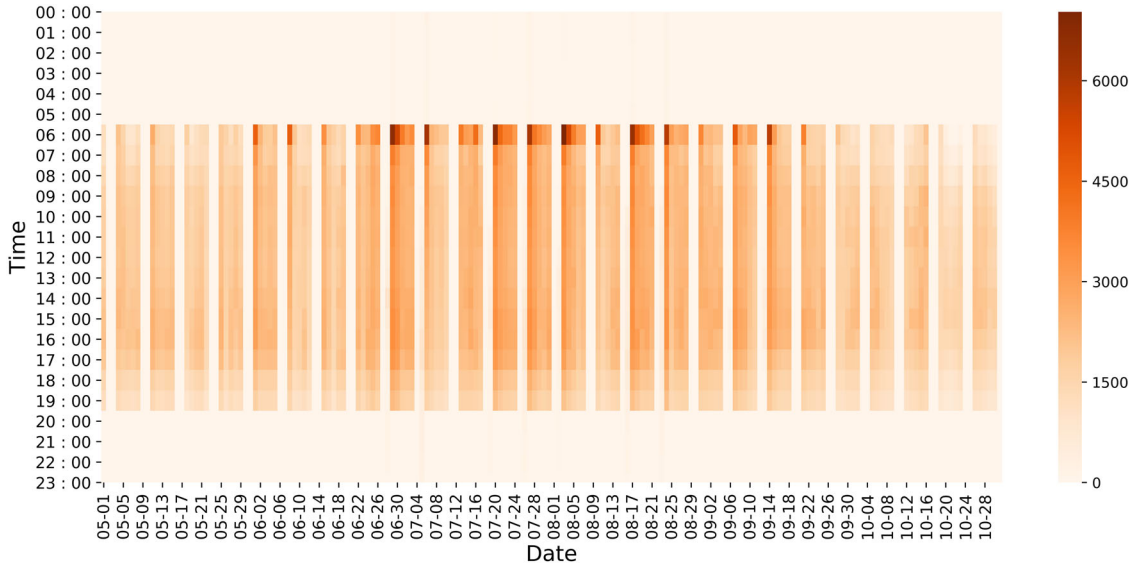

Fig. 5. System cooling load data simulated by EnergyPlus (hourly, kW).

In Figure 5, the color of a cell indicates the system total cooling load at a certain time. The peak load of the system appears typically in the morning, in particular on Monday mornings. In addition, there is no cooling load on weekends.

Simulation platform and system model

The hourly simulation case study is conducted with Python. In this study, Equations 6–8 comprise the chiller model, and Equation 6 is the standard chiller model adopted in EnergyPlus and DOE-2.1 (DOE-22, 2020), named "Electric Chiller Model-Based on Condenser Entering Temperature." Equations 9 and 10 comprise the pump model, and Equations 9–11 constitute the cooling tower model developed by CoolTool (Benton et al. 1999), which is also adopted by EnergyPlus. As Figure 6 shows, in every simulation step, the variable values are circulated in these three equipment models until T_{cwr} converges.

Equation 11 is a simplified version of the CoolTool model with several fixed variables (airflow rate ratio and water flow rate ratio) because the case cooling towers have constant speeds. Moreover, Equations 9 and 10 are simplified equations for calculating the power and flow rate because the pumps and cooling tower fans have constant speeds in this study.

$$\begin{cases} ChillerCapFTemp = b_1 + b_2 T_{chws} + b_3 {T_{chws}}^2 + b_4 T_{cwr} \\ \qquad\qquad + b_5 {T_{cwr}}^2 + b_6 T_{cwr} T_{chws} \\ ChillerEIRFTemp = d_1 + d_2 T_{chws} + d_3 {T_{chws}}^2 + d_4 T_{cwr} \\ \qquad\qquad + d_5 {T_{cwr}}^2 + d_6 T_{cwr} T_{chws} \\ ChillerEIRFPLR \;\; = \;\; g_1 + g_2 PLR + g_3 PLR^2 \\ P_{chiller} = P_{ref} \times (ChillerCapFTemp) \times (ChillerEIRFTemp) \\ \qquad \times (ChillerEIRFPLR) \end{cases} \tag{6}$$

$$\begin{cases} T_{chws} = max\left[T_{chws,set}, T'_{chwr} - CC/\left(\dfrac{C_p \times F_{chw} \times \rho}{3600s/h}\right)\right] \\ T_{chwr} = T_{chws} + CL_c/\left(\dfrac{C_p \times F_{chw} \times \rho}{3600s/h}\right) \end{cases} \tag{7}$$

$$T_{cws} = T_{cwr} + (P_{chiller} + CL_c) \div \left(\frac{C_p \times F_{cw} \times \rho}{3600s/h}\right) \tag{8}$$

$$P = \begin{cases} P_{ref} \;\; if \;\; operating \\ 0 \;\; not \;\; operating \end{cases} \tag{9}$$

$$F = \begin{cases} F_{nominal} \;\; if \;\; operating \\ 0 \;\; not \;\; operating \end{cases} \tag{10}$$

$$\begin{aligned} approach = {} & k_1 + k_2 T_{wet} + k_3 {T_{wet}}^2 + k_4 {T_{wet}}^3 + k_5 \Delta T_{cw} \\ & + k_6 {\Delta T_{cw}}^2 + k_7 {\Delta T_{cw}}^3 \end{aligned} \tag{11}$$

where T_{chws} is the temperature of the chilled water leaving the chiller (°C), T_{cwr} the temperature of the cooling water entering the chiller (°C), and P_{ref} the nominal power (kW); T_{chwr} is the temperature of the chilled water returning to the chiller (°C), T'_{chwr} is the T_{chwr} of the last time step, CC is the nominal cooling capacity of the chiller (kW), F_{chw} is the rate of chilled water flowing through the chiller evaporator (m^3/h), C_p is the specific heat capacity of water = 4.2 kJ/(kg·K), ρ is the water density = 1000 kg/m^3, and CL_c is the cooling load acting on the chiller (kW); T_{cws} is the temperature of the cooling water leaving the chiller (°C), F_{cw} is the cooling water flow rate through the chiller condenser (m^3/h), $F_{nominal}$ is the nominal flow rate of a pump or fan (m^3/h), T_{wet} is the ambient wet bulb temperature (°C), ΔT_{cw} is the difference between T_{cws} and T_{cwr}, and "approaches" the difference between T_{cwr} and T_{wet}; and finally, b_i, d_i, g_i, and k_i are coefficients determined by the engineering reference and datasets provided by EnergyPlus (these coefficients are

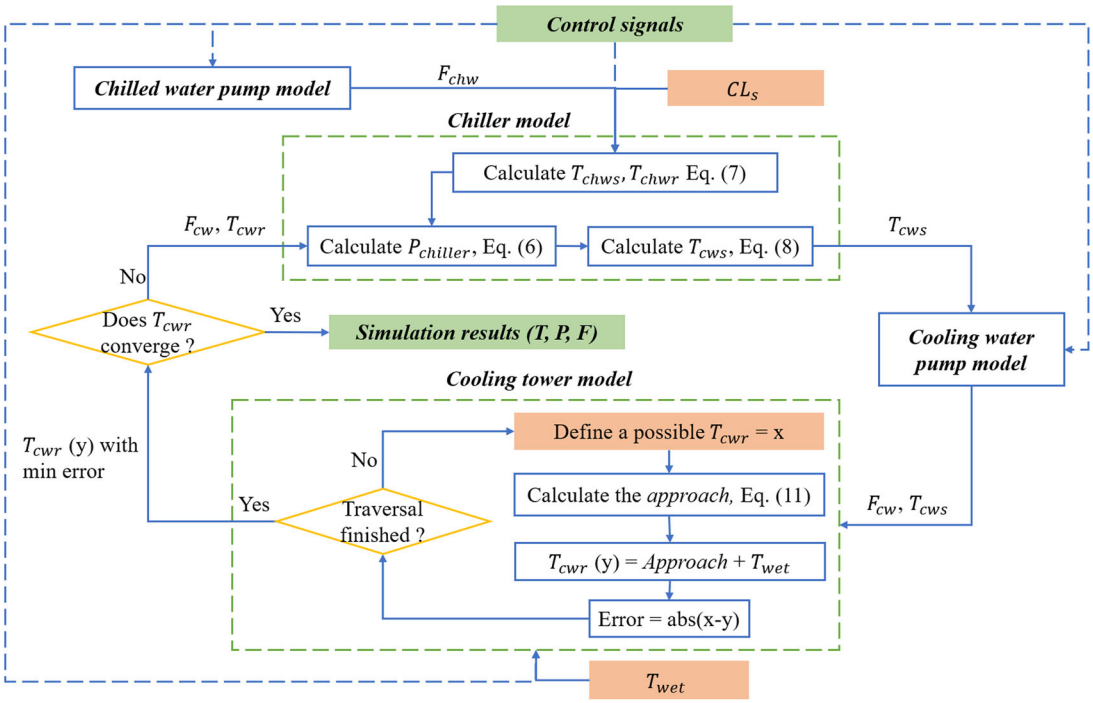

Fig. 6. Simulation process in one simulation step.

addressed in Table A1 and A2 of Appendix A of this article).

The simulation process is illustrated in Figure 6. When a time step begins, T_{chws}, T_{chwr}, F_{chw}, F_{cw}, P_{tower}, P_{chwp}, and P_{cwp} are calculated first without traversal or iteration; T_{cwr} and T_{cws} are two coupled variables because T_{cwr} influences the chiller performance, which in turns influences T_{cws}; the latter influences the outlet water temperature of the cooling tower (T_{cwr}). Hence, these two variables require iteration between the chiller model and cooling tower model to determine them (the big loop from model to model; Figure 6). Furthermore, in the cooling tower model, T_{cwr} requires traversal to be determined. The process is as follows: Define the T_{cwr} value (x); calculate the approach; calculate the corresponding T_{cwr} value (y) and the error between x and y; traverse each possible T_{cwr} value (x) to determine the (x, y) with the minimal error. In the cooling tower model, T_{cwr} is determined with traversal instead of iteration because the convergence performance of this cooling tower model is unacceptable (diverges easily).

The previously presented system model is established only for the simulation study and not embedded in the proposed model-free method. Only the information specified in overview section is embedded in the model-free controller beforehand.

Realization details of proposed method and two comparative loading methods

In this study, three loading methods including the proposed one are used to validate the performance of the proposed method. The sequencing control process of these controllers are the same as in the subsection on sequencing control of chiller plants. To protect the hardware and maintain indoor comfort, $T_{chws,set}$ remains within 5–9 °C. The two comparative methods are described next:

(1) *Baseline loading:* $T_{chws,set}$ is set to 6°C (nominal value) for all operating chillers.

(2) *Model-based loading:* The Lagrange multiplier (Bertsekas 1982)-based OCL method developed by Chang (Chang 2004) is adopted in this study as the compared model-based method, and Equation 12 is the auxiliary function. The optimization objective of this method is the minimal total power of the chillers (first term in Equation 12). The constraints are the second term in Equation 12. Moreover, the calculated PLRs are limited to 0.4–1, such as in the model-free method. The accurate chiller performance models (Equation 11 with accurate coefficients in Table A1) are embedded in the model-based controller before the simulated operation. In every time step, the model-based controller first calculates λ with Equation 13 and determines the PLRs of each running chiller with Equation 14; and finally, $T_{chws,set}$ is set with Equation 5 for each running chiller:

$$ChillerEIRFPLR_j = g_{1,j} + g_{2,j} \times PLR_j + g_{3,j} \times PLR_j^2 \tag{11}$$

$$L = \sum_{j=1}^{m} (ChillerEIRFPLR_j \times P_{ref,j}) + \lambda\left[CL - \sum_{j=1}^{m} (PLR_j \times CC_j)\right] \tag{12}$$

$$\lambda = \frac{2CL + \sum_{j=1}^{m} (CC_j \times g_{2,j} \div g_{3,j})}{\sum_{j=1}^{m} \frac{CC_j^2}{g_{3,j} \times P_{ref,j}}} \tag{13}$$

$$PLR_j = \frac{\lambda \times CC_j \div P_{ref,j} - g_{2,j}}{2g_{3,j}} \quad (14)$$

Regarding the Q-table in the case study, the state space is specified according to the maximal system cooling capacity. The initial values of the Q-table are set to 6, which is slightly higher than the nominal comprehensive chiller COP (Table 4).

Table 4. Initial Q-table in case study.

A/S	369 kW	738 kW	……	7380 kW
0	6	6		6
0.05	6	6		6
……				
1	6	6	……	6

Results and discussion

The operation of the case system from May 1 to October 31 (entire cooling season) is simulated on an hourly basis under the supervision of three loading methods. The simulation results are discussed based on two aspects: (1) chiller loading and sequencing, and (2) energy consumption and chiller COP.

Chiller loading and sequencing

The sequencing results of the case system are presented in Figure 7. The color of a cell in Figure 7 indicates the total cooling capacity of the system at a certain time. At most times during May, June, September, and October, the system cooling load is small and acts basically only on one chiller. In July and August, multiple chillers operate because the system cooling load frequently exceeds the cooling capacity of the big chiller, which allows the loading strategies to optimize the chiller loading. The operation time of chiller 2 is short because it is usually used to cope with the peak load, which appears occasionally.

Figure 8 presents each chiller load during the 6-month system operation. The scatter colors correspond to the different chillers. Figure 9 shows the PLR distribution of each chiller, and Figure 10 presents the relationship between the system cooling load and the ratio of system cooling load distributed to the small chiller.

(1) *Chiller loading results under baseline loading:* The left side of Figure 8 indicates that each chiller load is approximately linearly related to the system cooling load because the baseline loading sets $T_{chws,set}$ at 6°C (nominal value) for all operating chillers; consequently the ΔT_{chw} values of all operating chillers are equal. Hence, the cooling load on one chiller could be calculated by ($CL_j = F_{chw,j} / \sum F_{chw,j} \times CL_s$). The left side of Figure 9 suggests that compared to the other two OCL methods, the baseline loading method fails to enhance the PLR of the big chiller 1. This is one of the reasons why the COP of chiller 1 is lower under the baseline loading than under the other two loading methods (Figure A1 in Appendix A presents the positive relationship between PLR and COP of chillers 1 and 2).

(2) *Results of model-free loading method:* The distribution of the chiller loading results in Figure 8 is not very clear because the model-free controller must explore potential solutions and learn from the environmental reward (i.e., comprehensive chiller COP) to build its experience. The exploration procedure results in the irregular distribution of scatters. The stabilized loading actions of the model-free controller are illustrated in Figure 10. After learning for an entire cooling season, the loading action selected by the model-free controller is close to the optimal solution determined by the model-based controller. Moreover, Figure 10 indicates that both model-free loading and model-based loading tend to distribute more cooling load to the big chiller(s) compared to the baseline. That could enhance the comprehensive chiller COP because the peak COP of the big chillers occurs when the PLR is about 70% (Figure A1 in Appendix A), while the peak COP of the small chiller occurs at low PLR (Figure A1 in Appendix A).

(3) *Results of model-based loading method*: The right side of Figure 8 presents two inflection points (marked by the red circle). Before the points, the model-based controller tends to distribute less cooling load to the small chiller, more cooling load to the big chiller(s), because the peak COP of the small chiller occurs at low PLR (Figure A1 in Appendix A), contrary to how this takes place in big chillers; afterward, the trade-off between the two differently sized chillers is reevaluated, and more cooling load is assigned to the small chiller.

(4) The points marked by the black circle represent the first operation step, in which all chillers are switched on to cool the grid water.

Energy consumption and chiller COP

The simulated energy consumptions of the case system with the three loading methods are listed in Table 5. The energy consumed by the pumps and cooling towers does not change with the loading method because the three methods adopt the same sequencing control logic for controlling the chiller plants. However, the chiller energy is related to the loading method.

Baseline loading results in the highest chiller energy because this method does not optimize the cooling load or T_{chws} for the chillers. Compared to the baseline, the model-free loading method saves approximately 4.36% of chiller energy, which is less than the 4.95% of the model-based method. This is because the model-free method is not applied with a priori knowledge of the chiller performance, unlike the model-based method, which can accurately solve the conditional extremum (Bertsekas 1982).

Figure 11 shows the daily energy conservation of two optimal loading methods compared to that of the baseline. Although the model-free method is not embedded with the chiller models, it still saves chiller energy from the

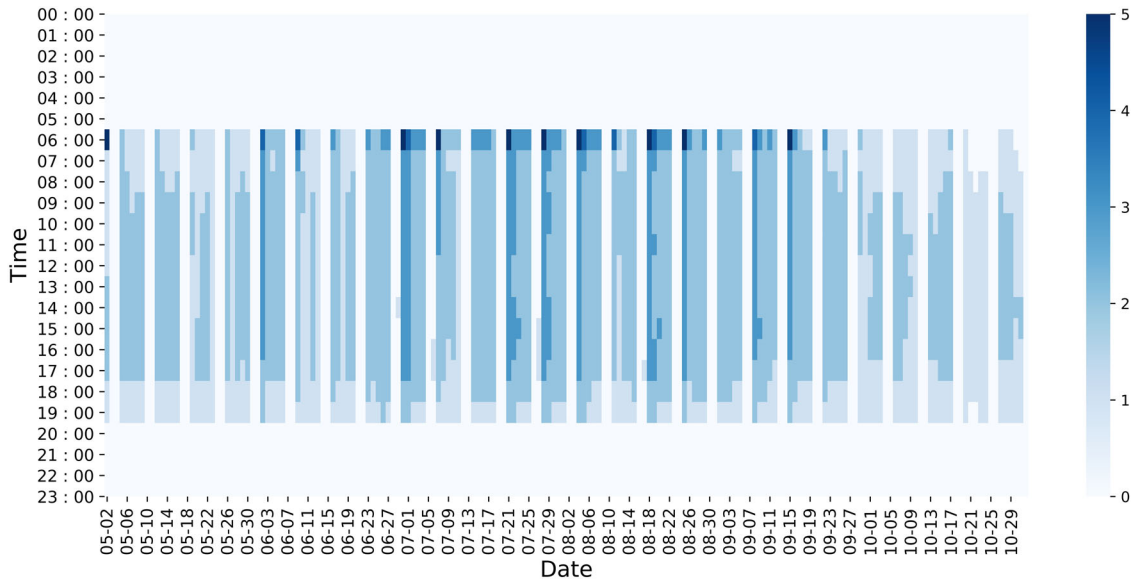

Fig. 7. Chiller sequencing result. (0: System off, 1: Chiller 3# operating, 2: Chiller 1# operating, 3: Chillers 1# and 3# operating, 4: Chillers 1# and 2# operating, 5: All chillers operating).

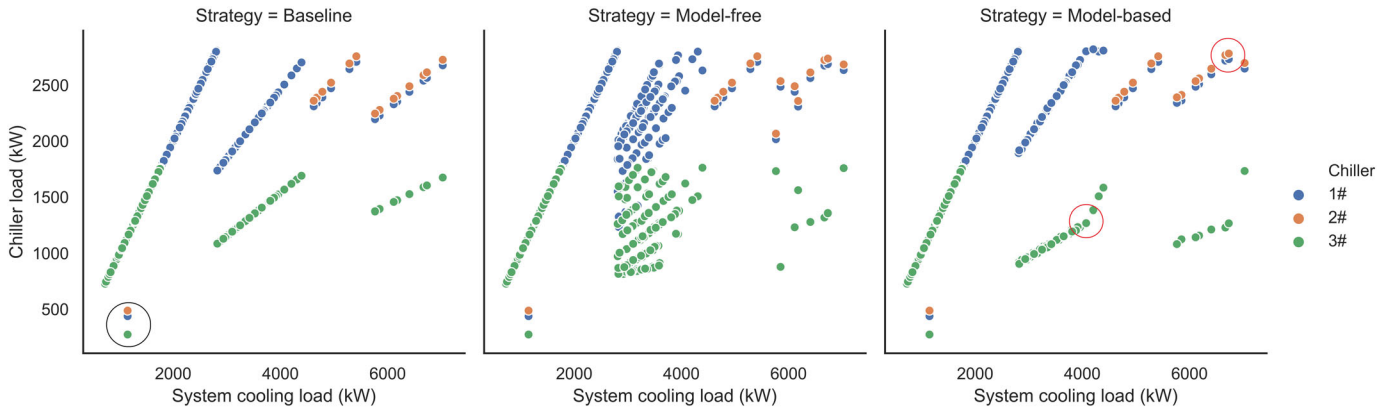

Fig. 8. Chiller loading results: chiller load versus system cooling load.

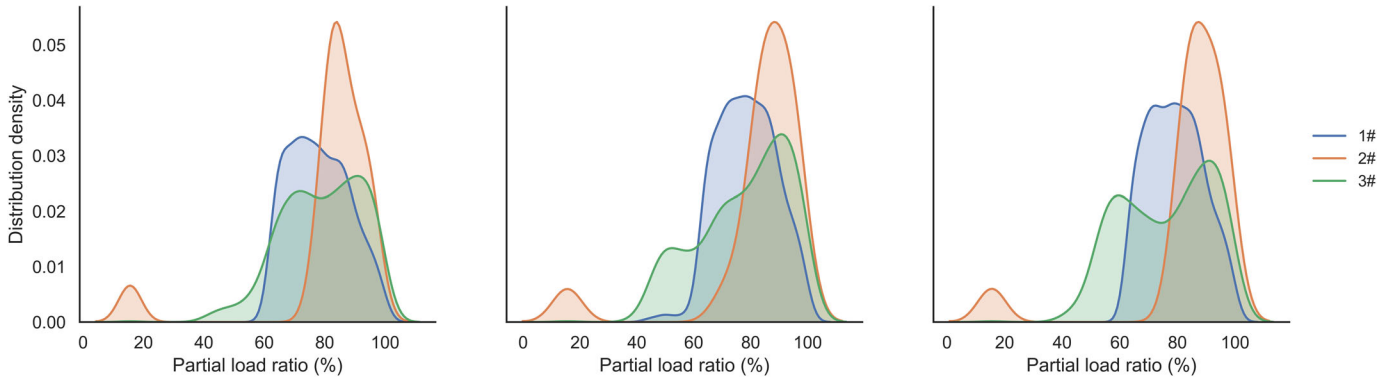

Fig. 9. Chiller loading results: distribution of PLR of each chiller.

beginning. This is because the optimization logic of T_{chws} (i.e., Equation 5) is adopted in the model-free loading method to enhance the chiller COP (Figures 12 and 13), which is not considered in the baseline control.

The scatters in Figure 11 suggest that more energy is conserved in the middle of the cooling season (July and August). This can be explained with the system cooling load profile (Figure 5) and chiller sequencing result (Figure 7): (1) In the middle of the cooling season, the hourly system cooling load is relatively high, which results in a higher potential for adjusting the chiller load; and (2) when the system cooling load is smaller than the capacity of the big chiller, only one chiller

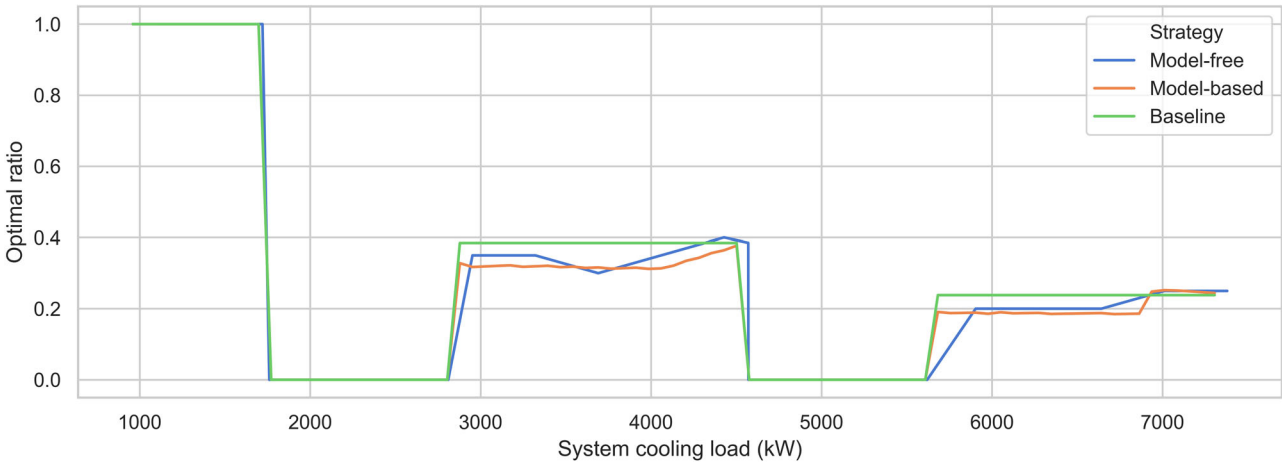

Fig. 10. State (system cooling load) versus optimized loading action (ratio of system cooling load acting on small chiller) under different loading methods.

Table 5. Energy consumption of case system with three loading methods.

	Cooling tower energy (kWh)	Cooling water pump energy (kWh)	Primary chilled water pump energy (kWh)	Chiller energy (kWh)	System energy consumption (kWh)
Baseline	50 779.5	136 890.0	49 695.0	712 476.0	950 110.5
Model-free loading				681 432.1	918 796.6
Model-based loading				677 221.4	914 585.9

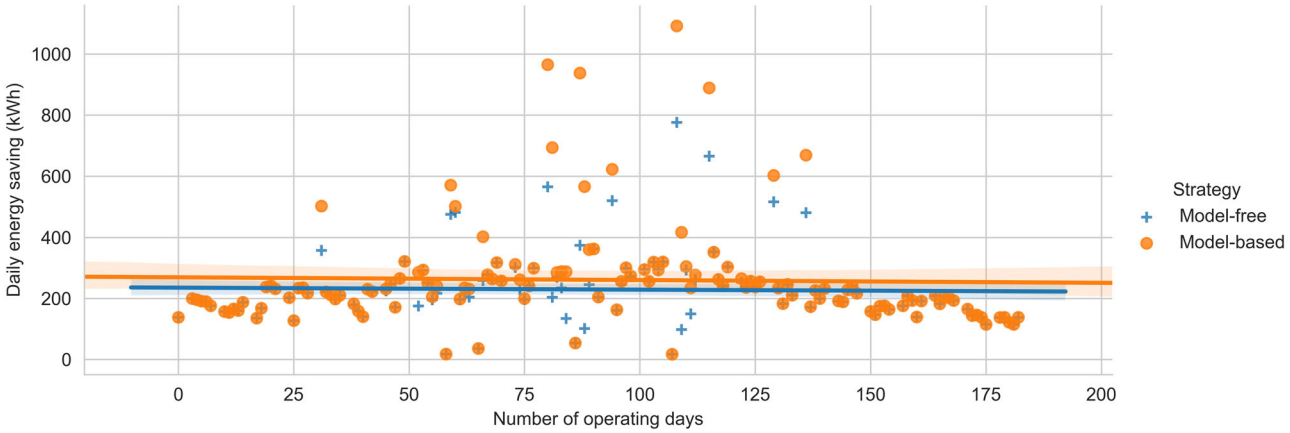

Fig. 11. Daily energy saving with model-free and model-based methods.

must operate (one big or one small), which means that no loading occurs; at this time, the energy-saving performance of the OCL methods is not exploited.

The COP simulation results of each chiller are presented in Figure 14:

(1) The COPs of chillers 1 and 3 are higher with the two optimal loading methods compared to those under baseline loading; this verifies the energy-saving effect of the two optimal loading methods. The COP of chiller 2 does not change much with the loading methods because chiller 2 is mainly used to cope with the peak load; in addition, there is not sufficient latitude for optimizing its loading.

(2) The COP outliers of chiller 2 are equal for all three loading methods. They result from the first step of the system operation when all chillers are switched on to cool the grid water (black circle in Figure 8).

(3) The differences between the chiller COPs of the three loading methods are not evident because most hourly cooling load values of the case system are not sufficiently high for the OCL, which limits the performance of the two optimal loading methods.

Evolution of the energy saving performance

The training time of an RL algorithm is important to its application (Sutton, Barto, and Bach 2018). Five rounds of

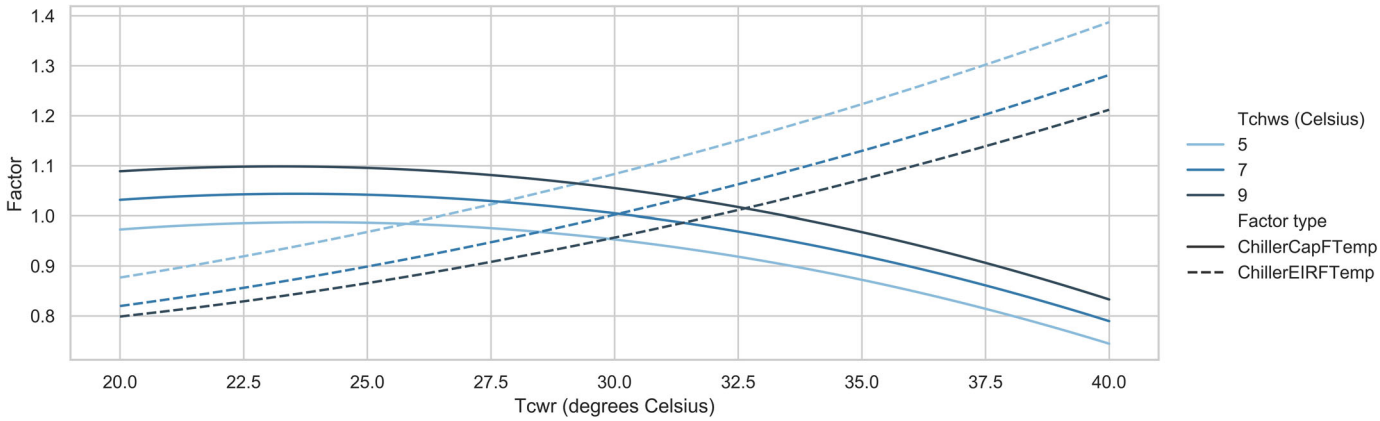

Fig. 12. Performance curves of big chillers: ChillerEIRFTemp and ChillerCapFTemp.

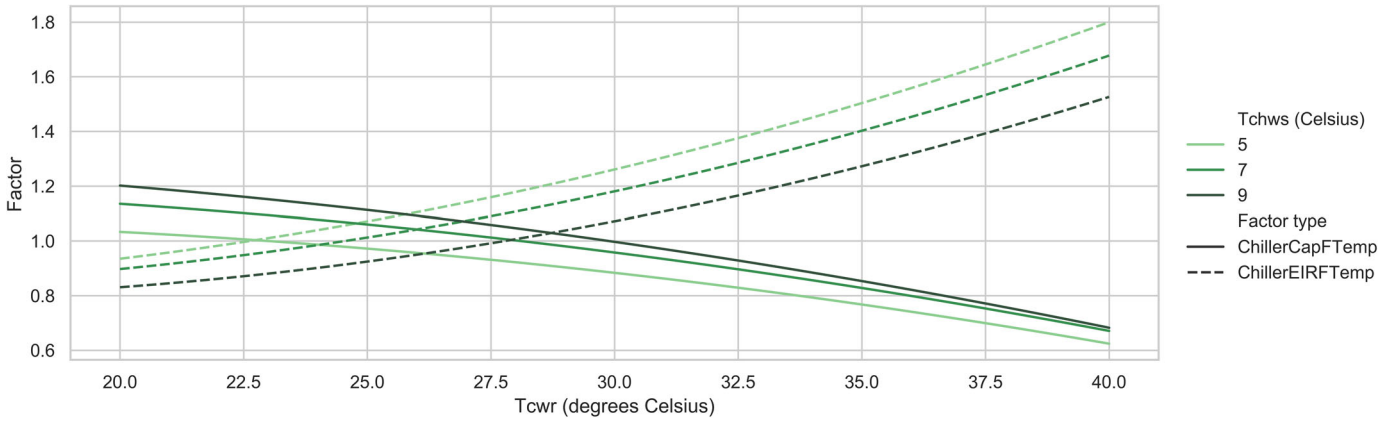

Fig. 13. Performance curves of small chiller: ChillerEIRFTemp and ChillerCapFTemp.

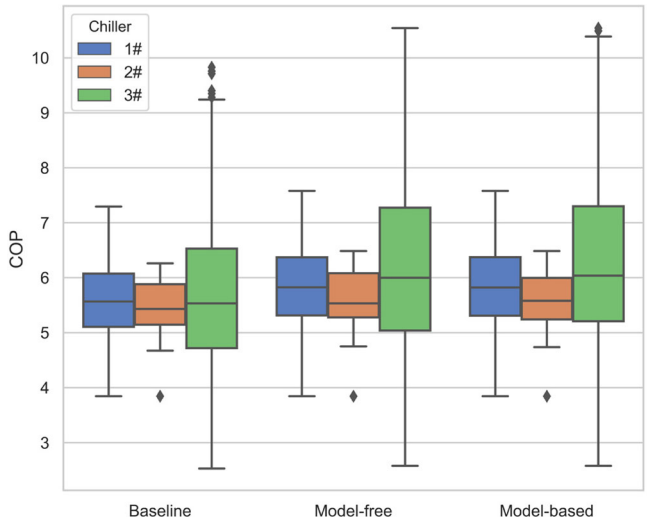

Fig. 14. Distributions of chiller COPs of three loading methods.

simulation are conducted to investigate the evolution of the proposed loading method. In each round, the model-free controller functions and evolves continuously for 10 years (10 cooling seasons). The simulation of each year adopts the same weather data and cooling load data introduced in the section on the case system . The results of chiller energy saving are listed in Table 6. Average chiller energy saving rates compared to the baseline are illustrated in Figure 15.

Table 6 indicates that the evolution procedure of the model-free controller in each round is different. That is because the training period has certain randomness due to the stochastic optimization policy (Equation 4).

Table 6. Annual chiller energy saving during 10 years (10 cooling seasons).

	Annual chiller energy saving (kWh)						
Year	Model-free round 1	Model-free round 2	Model-free round 3	Model-free round 4	Model-free round 5	Model-free average	Model-based
1	30 493.2	31 005.7	31 151.3	30 812.8	31 756.6	31 043.9	35 254.6
2	31 000.4	32 859.1	31 039.3	32 283.3	31 580.8	31 752.6	35 254.6
3	32 910.5	31 596.3	32 173.6	32 867.2	32 345.8	32 378.7	35 254.6
4	31 960.8	32 684.0	33 477.3	33 315.9	32 718.6	32 831.3	35 254.6
5	32 557.2	32 796.8	32 915.4	33 170.0	32 199.5	32 727.8	35 254.6
6	32 662.4	33 389.3	33 358.5	33 294.4	31 444.7	32 829.9	35 254.6
7	33 337.6	33 308.9	32 932.5	33 399.6	32 710.3	33 137.8	35 254.6
8	32 910.9	34 089.0	33 187.2	33 707.8	32 716.7	33 322.3	35 254.6
9	33 672.6	33 172.4	33 904.3	32 769.5	32 637.0	33 231.2	35 254.6
10	33 486.5	32 654.7	33 884.7	33 870.0	33 045.3	33 388.2	35 254.6

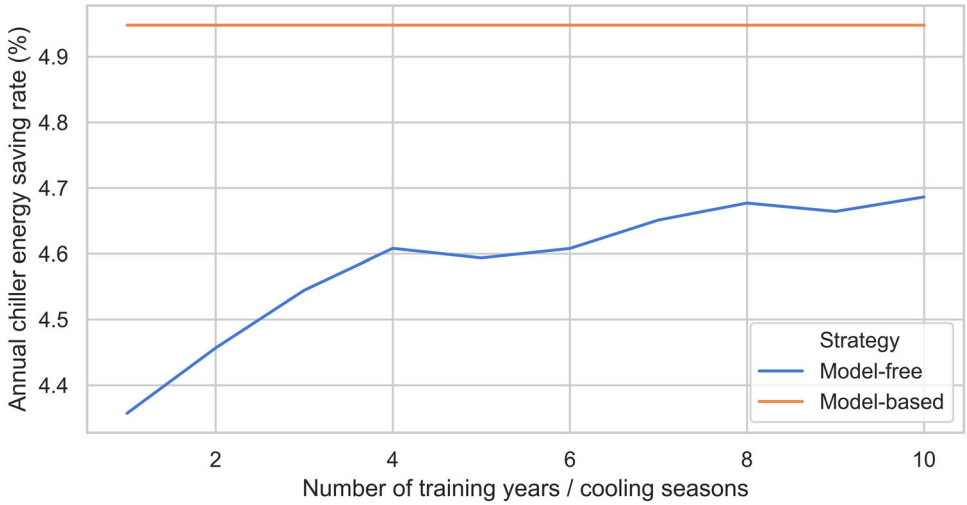

Fig. 15. Annual chiller energy saving rate.

Figure 15 shows that the energy-saving rate of the model-free controller increases to 4.6% in the first 4 years, and the performance is basically stable afterward. Although it takes 4 years to stabilize, the learning in the first year is the most important and evident. The energy-saving rate in the first year (4.36%) is already acceptable for application compared to the 4.95% of the model-based controller.

Conclusion and future work

Conclusion

An RL-based model-free control method for optimizing the chiller loading for central chiller plants is proposed. In this method, the comprehensive chiller COP is the optimization objective, and $T_{chws,set}$ is the control signal for realizing the loading solution. The discretized system cooling load is considered a state. The system layout and equipment name plates are used as a priori knowledge. Moreover, a simulation case study based on standard equipment models of EnergyPlus is conducted on the Python platform, under the supervision of three chiller loading methods: baseline, model-based, and the proposed model-free loading methods. The 6-month (entire cooling season) simulation results indicate that:

(1) Compared with the baseline loading, the model-free loading method can conserve 4.36% chiller energy in the first cooling season, which is inferior to that of the model-based method (4.95%).
(2) Without a priori knowledge, the model-free controller is trained and used simultaneously during the 6-month system operation. At the end of the operation period, the model-free controller's loading decision is close to that of the model-based controller.
(3) The hourly optimization by the model-free controller takes less than 0.1 s, which is acceptable for engineering applications.
(4) The energy-saving performance of the model-free controller could be furtherly improved in a longer applied period. Although it takes 4 years (four cooling seasons) for the annual chiller energy saving rate to stabilize at 4.6–4.7%, the energy saving rate in the first year (4.36%) is already acceptable for practical application.

Future work

The proposed model-free optimal chiller loading method is designed for the central chiller plant composed with two sizes of chillers (big chiller(s) and small chiller(s)). Only one RL agent is defined, for which the action is the ratio of the system cooling load assigned to small chillers.

(1) When more than two chiller sizes are used, the agent action space should be redesigned.
(2) Also, the proposed method is based on Q-learning, a table-based method, for which the state space and action space have to be limited and predefined. This would sacrifice the flexibility and precision of control. If more flexibility and higher control precision are required, a network-based RL method (e.g., deep Q network) may be more appropriate.
(3) Moreover, when the central chiller plant is complex, a multiple-agent RL can be used as a potential algorithm to optimize the system operation: for example, one agent to coordinate pumps, one to control chillers, one to control cooling towers, and so on.

Nomenclature

Cooling water system variables

CL_s = System cooling load (kW)
CL_c = Cooling load on the chiller (kW)
T_{chws} = Chilled water temperature (outlet of chillers) (°C)
T_{chwr} = Chilled water temperature (inlet of chillers) (°C)
T_{cwr} = Cooling water temperature (inlet of chillers) (°C)
T_{cws} = Cooling water temperature (outlet of chillers) (°C)
F_{chw} = Chilled water flow rate (m^3/h)
F_{cw} = Cooling water flow rate ($\frac{m^3}{h}$)
C_p = Specific thermal capacity of water ($kJ \cdot kg^{-1} \cdot K^{-1}$)
$T_{chws,set}$ = T_{chws} set point of chiller (°C)
T_{air} = Ambient dry bulb temperature (°C)
T_{wet} = Ambient wet bulb temperature (°C)
f = Equipment operating frequency (Hz)
P = Electrical power (kW)

Reinforcement learning variables

s = Last state (environment) of the agent
s' = Current state of the agent
a = Last action (operating frequency) done by the agent
a' = Potential action that the agent may execute in the current state
r = Reward to the agent for last action in last state
Q = Q-value in Q-table

Abbreviations

COP = Coefficient of performance
CHWP = Chilled water pump
CWP = Cooling water pump
PLR = Partial load ratio
CC = Chiller cooling capacity kW)
CL = Cooling load (kW)
OCL = Optimal chiller loading
OCS = Optimal chiller sequencing
RL = Reinforcement learning

Funding

The study has been supported by the National Key R&D Program of China (grant no. 2017YFC0702900) and Ministry of Science and Technology of the People's Republic of China.

References

Abdi, H., and L. J. Williams. 2010. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (4): 433–459. doi:10.1002/wics.101

Ardakani, A. J., F. F. Ardakani, and S. H. Hosseinian. 2008. A novel approach for optimal chiller loading using particle swarm optimization. *Energy and Buildings* 40 (12):2177–87. doi:10.1016/j.enbuild.2008.06.010

Arthur, D., and S. Vassilvitskii. 2007. K-Means++: The Advantages of Careful Seeding. In Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, LO, January 7–9.

ASHRAE. 2000. Liquid chilling system. In *ASHRAE systems and equipment handbook*. Atlanta, GA: ASHRAE Learning Institution.

Beghi, A., L. Cecchinato, and M. Rampazzo. 2011. A multi-phase genetic algorithm for the efficient management of multi-chiller systems. *Energy Conversion and Management* 52 (3):1650–61. doi:10.1016/j.enconman.2010.10.028

Benton, D. J., C. F. Bowman, M. Hydeman, and P. Miller. 1999. An improved cooling tower algorithm for the CooiTools™ simulation model. In ASHRAE Annual Meeting, Atlanta, GA. http://www.taylor-engineering.com/Websites/taylorengineering/articles/ASHRAE_Symposium_AC-02-9-4_Cooling_Tower_Model.pdf.

Bertsekas, D. P. 1982. *Constrained optimization and Lagrange multiplier methods*. New York: Academic Press.

Chang, Y. C. 2004. A novel energy conservation method: Optimal chiller loading. *Electric Power Systems Research* 69 (2–3):221–6. doi:10.1016/j.epsr.2003.10.012

Chang, Y. C., F. A. Lin, and C. H. Lin. 2005. Optimal chiller sequencing by branch and bound method for saving energy. *Energy Conversion and Management* 46 (13–14):2158–72. doi:10.1016/j.enconman.2004.10.012

Chang, Y. C., J. K. Lin, and M. H. Chuang. 2005. Optimal chiller loading by genetic algorithm for reducing energy consumption. *Energy and Buildings* 37 (2):147–55. doi:10.1016/j.enbuild.2004.06.002

Chen, C. W., Y. C. Chang, W. T. Liao, and C. W. Lee. 2014. Application of genetic programming method combined with neural network in HVAC optimal operation. *Applied Mechanics and Materials* 548–549:1030–1034. doi:10.4028/www.scientific.net/AMM.548-549.1030

Coelho, LdS, C. E. Klein, S. L. Sabat, and V. C. Mariani. 2014. Optimal chiller loading for energy conservation using a new differential cuckoo search approach. *Energy* 75:237–43. doi:10.1016/j.energy.2014.07.060

Commercial Buildings Energy Consumption Survey (CBECS). 2012. 2012 CBECS Survey Data. https://www.eia.gov/consumption/commercial/data/2012/.

Crawley, D. B., L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, et al. 2001. EnergyPlus: Creating a new-generation building energy simulation program. *Energy and Buildings* 33 (4): 319–331. doi:10.1016/S0378-7788(00)00114-6

DOE-2. 2020. http://www.doe2.com/.

EnergyPlus. 2020. https://energyplus.net/.

Fulkerson, B. 1995. Machine learning, neural and statistical classification. *Technometrics* 37 (4):459. doi:10.1080/00401706.1995.10484383

Gao, D. C., S. Wang, K. Shan, and C. Yan. 2016. A system-level fault detection and diagnosis method for low delta-T syndrome in the complex HVAC systems. *Applied Energy* 164:1028–1038. doi:10.1016/j.apenergy.2015.02.025

Henze, G. P., and J. Schoenmann. 2003. Evaluation of reinforcement learning control for thermal energy storage systems. *HVAC&R Research* 9 (3):259–275. doi:10.1080/10789669.2003.10391069

Hou, J., P. Xu, X. Lu, Z. Pang, Y. Chu, and G. Huang. 2018. Implementation of expansion planning in existing district energy system: A case study in China. *Applied Energy* 211:269–81. doi:10.1016/j.apenergy.2017.10.118

Kreider, J. F. 1994. *Heating and cooling of buildings: design for efficiency*. New York: McGraw-Hill.

Lee, W., and L. Lin. 2009. Optimal chiller loading by particle swarm algorithm for reducing energy consumption. *Applied Thermal Engineering* 29 (8–9):1730–4. doi:10.1016/j.applthermaleng.2008.08.004

Li, W., P. Xu, X. Lu, H. Wang, and Z. Pang. 2016. Electricity demand response in China: Status, feasible market schemes and pilots. *Energy* 114:981–94. doi:10.1016/j.energy.2016.08.081

Li, Z., G. Huang, and Y. Sun. 2014. Stochastic chiller sequencing control. *Energy and Buildings* 84 (84):203–13. doi:10.1016/j.enbuild.2014.07.072

Liao, Y., Y. Sun, and G. Huang. 2015. Robustness analysis of chiller sequencing control. *Energy Conversion and Management* 103: 180–190. doi:10.1016/j.enconman.2015.06.060

Liu, S., and G. P. Henze. 2006a. Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory. Part 2: Results and analysis. *Energy and Buildings* 38 (2):148–161. doi:10.1016/j.enbuild.2005.06.001

Liu, S., and G. P. Henze. 2006b. Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory: Part 1. *Energy* and Buildings 38 (2): 142–147. doi:10.1016/j.enbuild.2005.06.002

Liu, Z., H. Tan, D. Luo, G. Yu, J. Li, and Z. Li. 2017. Optimal chiller sequencing control in an office building considering the variation of chiller maximum cooling capacity. *Energy and Buildings* 140: 430–442. doi:10.1016/j.enbuild.2017.01.082

Pang, Z., and Z. O'Neill. 2018. Uncertainty quantification and sensitivity analysis of the domestic hot water usage in hotels. *Applied Energy* 232:424–42. doi:10.1016/j.apenergy.2018.09.221

Pang, Z., P. Xu, Z. O'Neill, J. Gu, S. Qiu, X. Lu, and X. Li. 2018. Application of mobile positioning occupancy data for building energy simulation: An engineering case study. *Building and Environment* 141:1–15. doi:10.1016/j.buildenv.2018.05.030

Pang, Z., Z. O'Neill, Y. Li, and F. Niu. 2020. The role of sensitivity analysis in the building performance analysis: A critical review. *Energy and Buildings* 209:109659. doi:10.1016/j.enbuild.2019.109659

Pérez-Lombard, L., J. Ortiz, and C. Pout. 2008. A review on buildings energy consumption information. *Energy and Buildings* 40 (3): 394–8. doi:10.1016/j.enbuild.2007.03.007

Qiu, S., F. Feng, W. Zhang, Z. Li, and Z. Li. 2019. Stochastic optimized chiller operation strategy based on multi-objective optimization considering measurement uncertainty. *Energy and Buildings* 195:149–60. doi:10.1016/j.enbuild.2019.05.006

Qiu, S., F. Feng, Z. Li, G. Yang, P. Xu, and Z. Li. 2019. Data mining based framework to identify rule based operation strategies for buildings with power metering system. *Building Simulation* 12 (2):195–205. doi:10.1007/s12273-018-0472-6

Qiu, S., Z. Li, Z. Pang, W. Zhang, and Z. Li. 2018. A quick auto-calibration approach based on normative energy models. *Energy and Buildings* 172:35–46. doi:10.1016/j.enbuild.2018.04.053

Sulaiman, M. H., M. I. M. Rashid, M. R. Mohamed, O. Aliman, and H. Daniyal. 2015. An application of cuckoo search algorithm for solving optimal chiller loading problem for energy conservation. *Applied Mechanics and Materials* 793:500–504. doi:10.4028/www.scientific.net/AMM.793.500

Sun, Y., S. Wang, and G. Huang. 2009. Chiller sequencing control with enhanced robustness for energy efficient operation. *Energy and Buildings* 41 (11):1246–55. doi:10.1016/j.enbuild.2009.07.023

Sutton, R. S., A. G. Barto, and F. Bach. 2018. *Reinforcement Learning: An Introduction*. London: Bradford Book.

Swider, D. J. 2003. A comparison of empirically based steady-state models for vapor-compression liquid chillers. *Applied Thermal Engineering* 23 (5):539–56. doi:10.1016/S1359-4311(02)00242-9

Taylor, S. T. 2012. Optimizing design and control of chiller plants – Part 5: Optimized control sequences. *ASHRAE Journal* 2012: 56–74.

Wang, S., and F. Xiao. 2004. AHU sensor fault diagnosis using principal component analysis method. *Energy and Buildings* 36 (2):147–160. doi:10.1016/j.enbuild.2003.10.002

Wang, S., and Z. Ma. 2008. Supervisory and optimal control of building HVAC systems: A review. *HVAC&R Research* 14 (1): 3–32. doi:10.1080/10789669.2008.10390991

Wang, Y., X. Jin, Z. Du, and X. Zhu. 2018. Evaluation of operation performance of a multi-chiller system using a data-based chiller model. *Energy and Buildings* 172:1–9. doi:10.1016/j.enbuild.2018.04.046

Yang, L., Z. Nagy, P. Goffin, and A. Schlueter. 2015. Reinforcement learning for optimal control of low exergy buildings. *Applied Energy* 156:577–586. doi:10.1016/j.apenergy.2015.07.050

Zhu, D., T. Hong, D. Yan, and C. Wang. 2013. A detailed loads comparison of three building energy modeling programs: EnergyPlus, DeST and DOE-2.1E. *Building Simulation* 6 (3): 323–335. doi:10.1007/s12273-013-0126-7

Zhu, N., K. Shan, S. Wang, and Y. Sun. 2013. An optimal control strategy with enhanced robustness for air-conditioning systems considering model and measurement uncertainties. *Energy and Buildings* 67 (4):540–50. doi:10.1016/j.enbuild.2013.08.050

Appendix A

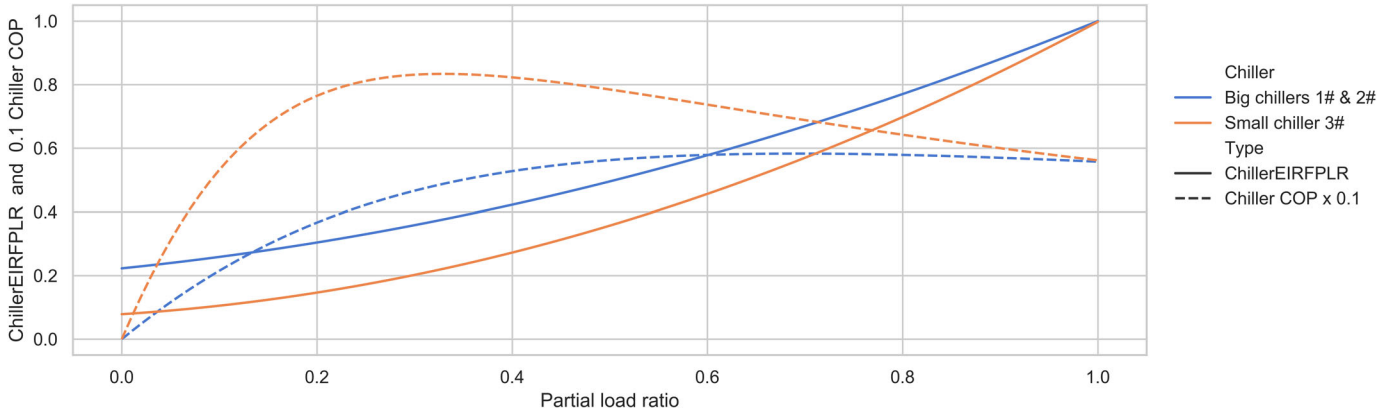

Fig. A1. PLR–ChillerEIRFPLR and PLR–COP curves.

Table A1. Details of chiller models adopted in the case study (according to EnergyPlus V8.0 Engineering reference.pdf, and Datasets/Chillers.idf).

	Big chillers 1 and 2 DOE-2.1 generic centrifugal chiller (idf row no. 11170)			Small chiller 3 York VSD centrifugal chiller (idf row no. 5684)		
Coefficients	*ChillerCapFTemp*	*ChillerEIRFTemp*	*ChillerEIRFPLR*	*ChillerCapFTemp*	*ChillerEIRFTemp*	*ChillerEIRFPLR*
b_1, d_1, g_1	0.257896	0.933884	0.222903	0.4575085	0.6794525	0.07859908
b_2, d_2, g_2	0.0389016	−0.058212	0.313387	0.1313508	0.06694756	0.1950291
b_3, d_3, g_3	−0.00021708	0.00450036	0.463710	−0.004408831	−0.003625396	0.7241581
b_4, d_4	0.0468684	0.00243		0.01930354	−0.01018762	
b_5, d_5	−0.00094284	0.000486		−0.0005479641	0.001066394	
b_6, d_6	−0.00034344	−0.001215		−0.001376580	−0.002113402	

Table A2. Details of adopted cooling tower model (from EnergyPlus V8.0 Engineering reference.pdf, pp. 783 and 784).

Simplified coefficients in this study	Corresponding detailed coefficients in the reference	Value
k_1	Sum of Coeff(1)–Coeff(10)	0.66465498
k_2	Sum of Coeff(11)–Coeff(16)	−0.01630528
k_3	Sum of Coeff(17)–Coeff(19)	−0.00286450
k_4	Coeff(20)	0.00007604
k_5	Sum of Coeff(21)–Coeff(30)	2.19373107
k_6	Sum of Coeff(31)–Coeff(34)	−0.06708132
k_7	Coeff(35)	0.00080737